



Perfectionism and the Repugnant Conclusion

Simon Beard¹ 

Published online: 5 March 2019
© The Author(s) 2019

Abstract

The Repugnant Conclusion and its paradoxes pose a significant problem for outcome evaluation. Derek Parfit has suggested that we may be able to resolve this problem by accepting a view he calls ‘Perfectionism’, which gives lexically superior value to ‘the best things in life’. In this paper, I explore perfectionism and its potential to solve this problem. I argue that perfectionism provides neither a sufficient means of avoiding the Repugnant Conclusion nor a full explanation of its repugnance. This is because even lives that are ‘barely worth living’ may contain the best things in life if they also contain sufficient ‘bad things’, such as suffering or frustration. Therefore, perfectionism can only fully explain or avoid the Repugnant Conclusion if combined with other claims, such as that bad things have an asymmetrical value relative to many good things. This combined view faces the objection that any such asymmetry implies Parfit’s ‘Ridiculous Conclusion’. However, I argue that perfectionism itself faces very similar objections, and that these are question-begging against both views. Finally, I show how the combined view that I propose not only explains and avoids the Repugnant Conclusion but also allows us to escape many of its paradoxes as well.

According to

the Repugnant Conclusion: *Compared with the existence of many people who would all have some very high quality of life, there is some much larger number of people whose existence would be better, even though these people would all have lives that were barely worth living.*^{1,2}

¹ Derek Parfit, “Can we avoid the repugnant conclusion?,” *Theoria* vol. 82, no. 2 (2016), p. 110.

² This is not the only version of the Repugnant Conclusion. Many philosophers now engage with versions like that given by Gustaf Arrhenius: “For any perfectly equal population with very high positive welfare, there is a population with very low positive welfare which is better” (see Gustaf Arrhenius, “An impossibility theorem for welfarist axiologies,” *Economics & Philosophy* vol. 16, no. 2 (2000), p. 248). One difference between these versions of the Repugnant Conclusion is that Parfit’s refer to people’s

✉ Simon Beard
Sjb316@cam.ac.uk

¹ Centre for the Study of Existential Risk, University of Cambridge, Cambridge, UK

This conclusion, which is implied by Total Utilitarianism and other moral theories, is illustrated below - where A is the large population with very high-quality lives, and Z is the much larger population with lives that are barely worth living (Fig. 1):

Many people intuitively find this conclusion very hard to accept. However, this intuitive repulsion cannot, by itself explain why this conclusion is unacceptable, nor does it indicate how we should avoid it. In fact, there are many arguments for accepting the Repugnant Conclusion, not least of which being that it is implied by principles that people find equally hard to reject, such as:

- a. the 'simple view' that "Anyone's existence is in itself good if this person's life is worth living. Such goodness has non-diminishing value, so if there were more such people, the combined goodness of their existence would have no upper limit."³ Together with,
- b. the principle that the value of all lives can be represented on a finite common scale, so that, compared with the existence of any number of very high-quality lives, there is always some number of other 'good' lives whose existence would be better, regardless of their quality.⁴ This principle is not explicitly stated by Parfit, but we might refer to it as the 'principle of non-lexicality' or, following Arrhenius, the 'Archimedean Principle'.⁵

This contradiction between intuitively plausible principles and their intuitively repugnant conclusion has created significant difficulties in the evaluation of outcomes involving different sized populations, such as climate policies that might affect the size and welfare of future generations.⁶

Avoiding the Repugnant Conclusion, therefore, requires accepting a moral theory that gives us some reason to reject, at least, one of these two principles. Merely proposing a moral theory that does not imply the Repugnant conclusion is not a sufficient response to this problem; we also need reasons for accepting this theory, rather than any of those theories that imply the Repugnant Conclusion instead. Any

Footnote 2 (continued)

'Quality of Life', a value that, though closely related to welfare, is not the same thing. Since this paper primarily deals with Parfit's response to the Repugnant Conclusion, I will focus on his own version of that conclusion to the exclusion of others.

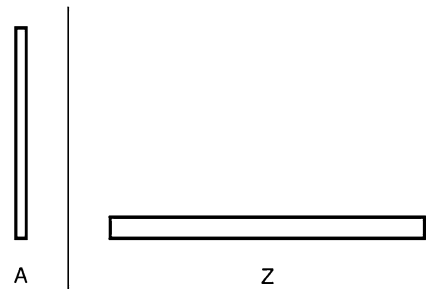
³ Ibid, p. 111.

⁴ See *ibid*, p. 112 and 114.

⁵ This is far from the only set of principles that imply the Repugnant Conclusion. Another set includes the 'mere addition' principle, that the addition of lives that are all worth living makes an outcome better, and the 'non-anti egalitarian' principle, that a greater quantity of welfare distributed equally across the same number of lives is better than a lesser quantity distributed unequally (Derek Parfit, *Reasons and persons* (Oxford: OUP, 1984), pp. 425–430). Together with the Repugnant Conclusion, these principles form the 'mere-addition paradox' that I discuss in section 5.

⁶ See Ottmar Edenhofer, Ramón Pichs-Madruga, Youba Sokona and Jan Christoph Minx, *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: CUP, 2014), p. 223 and Gustaf Arrhenius, *Population Ethics: The Challenge of Future Generations* (Oxford: OUP, forthcoming).

Fig. 1 The Repugnant Conclusion. The higher of each box represents people's Quality of Life, whilst its width represents the size of the population



solution to the problem of the Repugnant Conclusion must, therefore, provide both a means of avoiding the conclusion and a moral explanation of its repugnance.

In this paper, I will explore one proposed solution and the explanation it provides. This is that, whilst there is more of what makes lives worth living in the Z population than in the A population, there are qualitative differences between the kind of good things in the lives of these two populations that are more important than this quantitative difference. This would imply that we reject principle b, the principle of non-lexicality, because we must now evaluate lives on both a quantitative and a qualitative scale and may give lexical superiority to certain qualitative losses over any quantitative loss.

According to one view, proposed by Derek Parfit and known as perfectionism, this difference consists in the fact that high-quality lives not only contain much that makes these lives worth living but also some of ‘the best things in life’ – especially good things with a value that is lexically superior to that of other sources of welfare. The presence of these things in the A population, but not the Z population, constitutes our reason for rejecting the Repugnant Conclusion and moral theories that imply it. Whilst many have argued against perfectionism, philosophers usually assume that, if it were true, this view would both explain why we find the Repugnant Conclusion very hard to accept and allow us to avoid it.^{7,8}

⁷ See Derek Parfit, “Overpopulation and the Quality of Life,” in Peter Singer, ed. *Applied Ethics* (Oxford: OUP, 1986). Reprinted in Jasper Ryberg and Torbjörn Tännsjö, Eds., *The Repugnant Conclusion. Essays on Population Ethics* (Dordrecht: Kluwer Academic Publishers, 2004), James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: The Clarendon Press, 1986), p. 87, Roger Crisp, “Utilitarianism and the Life of Virtue,” *Philosophical Quarterly*, no. 42 (1992) pp. 150–151, Stewart Rachels, “Repugnance or Intransitivity: A Repugnant but Forced Choice,” In Jasper Ryberg and Torbjörn Tännsjö, Eds., *The Repugnant Conclusion. Essays on Population Ethics* (Dordrecht: Kluwer Academic Publishers, 2004) pp. 177–178, Michael Huemer, “In Defense of Repugnance,” *Mind* no. 468 (2008) pp. 914–915 and Parfit (2016) op. cit.

⁸ It is important to distinguish this use of the term perfectionism from others commonly found in moral and political philosophy. For instance, John Rawls uses the term to describe an approach to political philosophy that requires everybody to seek to maximize “achievement of human excellence in art, science and culture” (John Rawls, *A theory of justice* (Cambridge Ma: Harvard university press, 1971), 325). This differs from Parfit’s use of the term perfectionism in several ways, in particular in its failure also to value other things that make life worth living alongside these perfectionist goods. Other philosophers, such as Thomas Hurka, use the term perfectionism to refer to neo-Aristotelian accounts of the value of lives, on which value depends upon the achievement of one’s human potential rather than on feelings or desires. Naturally, fully achieving one’s potential could well constitute one of the best things in life; however, this form of perfectionism is a view about the nature of welfare in general, making it considerably broader than Parfit’s. Finally, and least relevant to this discussion, perfectionism has a political sense

1 The Nature of Perfectionism

Sadly, the details of perfectionism remain undeveloped. Derek Parfit, the view's main proponent, only briefly stated it twice in his published works. In his classic paper 'Overpopulation and the Quality of Life' he explained that according to Perfectionism "even if some change brings a great net benefit to those who are affected, it is a change for the worse if it involves the loss of one of the best things in life".⁹ More recently he stated that, on this view, if, for any population, "there would be no art, or science, no deep loves or friendships, no other achievements, such as that of bringing up our children well, and no morally good people: then that population could not be better than one in which all these things were present, even if, as with the Z and A populations, it contained much more welfare in total."^{10,11} In these statements, we are not to understand the terms 'welfare' or 'benefit' as referring to any particular good or action, such as fulfilling a person's desires or changing their mental states, but rather to increasing the quantity of whatever makes these lives worth living i.e. what promotes these people's self-interest by making their lives more valuable for them.¹²

There are two possible ways in which such claims might be justified, either by arguing that the best things in life make a person's life much better for them than other good things or by claiming that they make populations better in some other way.

On the first of these, the qualitative difference in the value of the best things in life compared to other good things that make lives worth living is a difference in their value for the people who enjoy them, which I will refer to as their welfare value. This is the classical way of making such claims. For instance, it seems to be what J.S. Mill intended by his famous claim, quoted approvingly by Parfit, that "we are justified in ascribing to [certain pleasures] a superiority in quality, so far outweighing quantity as to render it, in comparison, of small account".^{13,14} Defenders of such claims tend to point to the superior value of short periods containing

Footnote 8 (continued)

in which it is used to mean that society should not be neutral between different human aims, but should interfere in individual's lives if this will promote their real or higher interests.

⁹ Parfit (2004) op. cit. p. 19.

¹⁰ Parfit (2016) op. cit. p. 123.

¹¹ Despite proposing and doing much to defend this view, Parfit was keen to assert that he was not committed to it "I say that this view has some plausibility, but that it seems to me at times crazy. These remarks are not a way of accepting this view" (quoted from private correspondence). Nevertheless, it remained an important part of his proposed solution to the Repugnant Conclusion and he never explicitly mentioned any alternative view that might replace it. I hope that by expanding on and exploring this view, I can help in some small way to make it seem less crazy.

¹² It follows that a person's welfare level is an evaluation of the features of their life on its own, and does not depend upon other features of that person or the population in which they exist.

¹³ John Stewart Mill, *Utilitarianism*, Ed Roger Crisp (Oxford: OUP, 1998) p. 56.

¹⁴ See Parfit (1986) op. cit. p. 17. Although Parfit says that he agrees with this claim, he did so only once and, I shall argue, there are reasons to believe that this did not represent his final view on the subject.

high-quality welfare compared with any amount of lower quality welfare within a single life. For instance, Roger Crisp has argued that:

There is much more to [judgments of enjoyableness] than mere duration. There is nothing to prevent [somebody] claiming that it would not matter how long the experience of enjoyable drinking could be prolonged: she would never enjoy it as much as she [enjoys reading *Pride and Prejudice*]. For what she enjoyed in the novel was its wit, its beautiful syntax, and its exquisite delineation of character. The loss of such enjoyments ... – in the context of her own life – could never be compensated for, ... by any amount of lemonade enjoyment.¹⁵

Similarly, in the context of his objective goods account of welfare, Thomas Hurka writes:

However hard it is to accept a sacrifice of quality for quantity with pleasure or desire-fulfilment, it is even harder with goods of excellence, such as knowledge and achievement. (Think again of Achilles. Surely the loss of his greatest feats could not be made good by any number of successful shoelace-tyings.)¹⁶

Convincing as such arguments seem in the above passages, when evaluating different quantities and qualities of good things within individual lives, they cannot justify the perfectionist's claims about the value of the best things in life and the Repugnant Conclusion. One reason for this is that such arguments are far less convincing when we consider how to weigh the value of these things across different lives, especially when we consider the relative value of good and bad things in these lives. For instance, even if we accept that reading Jane Austin is more enjoyable than any amount of lemonade drinking within a single life, is this amount of enjoyment still greater than that of very many other people drinking lemonade (or of them all forgoing some equivalent pain such as a mild thirst)? Similarly, would we still consider the pursuit of human excellence to be served by Achilles retaining all his feats if this came at the cost of everybody losing the ability to tie their shoelaces? These accounts can only avoid the Repugnant Conclusion if the answer to such questions is yes, but I do not think it is.

Another, more important, reason why these arguments cannot justify the perfectionist's claims is that, when Parfit talked about the value of the best things in life, he clearly meant something more than their welfare value. If he had not, then the total value of the best things in life would already have been taken into account in assessing whether a particular change constituted a 'great net benefit' or whether one population had more welfare than another, yet his arguments require that this cannot be so.

This leads to the second way of justifying perfectionism's claims. On this view, the value of the best things in life consists not only of their welfare value (their value to the people who enjoy them) but also some other kind of value that lesser good

¹⁵ Roger Crisp, *Reasons and the Good* (Oxford, OUP, 2006) p. 115.

¹⁶ Thomas Hurka, *Perfectionism* (Oxford: OUP, 1993) p. 71.

things do not possess.¹⁷ Determining the nature of this ‘other’ value can seem difficult given that one of the most prominent features of discussions about welfare is the variety of claims about what welfare is.¹⁸ Yet, since we have already defined welfare as being anything that contributes to the value of a life for the person living it, it is safe to say that the extra value possessed by the best things in life must be some form of impersonal value.

The idea of lives having impersonal, as well as personal, value may at first strike us as odd – what can make a life better or worse that does not also make it better or worse for the person living it? Let alone make it so much better that a sufficient quantity of this good may be more valuable than any amount of welfare. However, there are several ways in which this may be true. Lives may, for instance, have an impersonal aesthetic value, either because of their own aesthetic qualities or because it is only when beautiful things are experienced within lives that their, impersonal, aesthetic value is fully realized (whether or not it actually contributes to the welfare of the person experiencing it). Alternatively, lives may have impersonal moral value, for instance because they are good lives in the sense of being worthy of ethical or moral approval or because they realize moral value that would otherwise not have been existed such the intrinsic value of being the subject of charity, love or care, which require that a person be benefited but whose full worth may exceed that of the benefit being received.

Two other promising views of how lives may come to have impersonal value are that they have a special value for a (temporal or psychological) part of a person, such as that person ‘as a philosopher’ or ‘as a child’, that does not extend to the person as a whole, or that they may have a special value for a group of people, such as a society, family or organization, that does not extend to its individual members. Whilst less commonly discussed, this kind of impersonal value can make sense if, like Parfit, one holds the view that ‘personhood’ is not a rigidly defined metaphysical category but rather a quality of having sufficient psychological continuity over time (such as shared beliefs, desires, intentions and traits), where one may also have weaker levels of psychological continuity with other people and may also have periods of much stronger psychological continuity within one’s life.

Some of these different kinds of impersonal value are captured in discussions about the value of ‘meaning’ in life, even though, as with welfare, there is no consensus about what such meaning is.¹⁹ Furthermore, many proponents of the value

¹⁷ I shall simply assume here that 1) the best things in life are the only goods that embody these additional values and 2) that the best things in life are also very good for the people who enjoy them. In personal correspondence, Derek Parfit suggested that instead of discussing different kinds of value it might be more appropriate to talk about the different ways in which goods can contribute to the overall value of an outcome.

¹⁸ See Derek Parfit, *Reasons and Persons* (Oxford: OUP, 1984) pp. 493–501 and Griffin op. cit.

¹⁹ Two characterisations of the value of meaning that have been especially well developed are: a) the “active engagement with projects of worth” (Susan Wolf, “Happiness and meaning: Two aspects of the good life,” *Social Philosophy and Policy* vol. 14, no. 1 (1997) p. 210) and b) a life’s, “largely in virtue of one’s actions and their causes and consequences, [warranting] great pride or admiration or [exhibiting] superlative final goods beyond one’s animal self.” (Thaddeus Metz, “The meaningful and the worthwhile: Clarifying the relationships,” *The Philosophical Forum*, vol. 43, no. 4 (2012), p. 447).

of meaning have argued that it is separate from and irreducible to the value of welfare.²⁰ Therefore, despite some differences in approach, I think that the value of meaning is probably the best existing candidate for the kind of impersonal value that is present in our enjoyment of the best things in life, but not other goods. Given this, together with the obvious connection between meaning and many of the goods listed by Parfit as constituting the best things in life, such as art, science, deep loves and friendships and other achievements, I shall use meaning as a proxy term for any impersonal value uniquely possessed by the best things in life, though I believe that the nature of such values is worthy of further study.²¹

If Parfit had in mind something like this second way of justifying perfectionism's claims, which I believe he must have, then the reasons we would have to reject the Repugnant Conclusion would be something like this. Even though the Z population has more welfare than the A population, the A population is valuable in ways that the Z population is not. The A lives are not only very good for the people who live them, they are also meaningful, or have some other kind of impersonal value, whilst the Z lives are not only much less good for the people living them but are also, because they lack the best things in life, devoid of these other values. Therefore, whilst the Z lives may together contain more welfare than the A lives, they cannot contain as much value, all things considered, because no amount of additional welfare can compensate for this loss.²²

This view faces many objections, and much of Parfit's writing on Perfectionism was aimed at addressing these. However, I am interested in whether it really constitutes an adequate response to the problem of the Repugnant Conclusion. In the next section, I will consider this question and argue that it does not. In Sect. 3 I will then propose an alternative view about the Quality of Life that combines Perfectionism with other principles and argue that this does allow us to produce a satisfactory response to the problem. I defend this combined view in Sects. 4 and 5.

²⁰ Susan Wolf, "Happiness and meaning: two aspects of the good life," *Social Philosophy and Policy* vol. 14, no. 1 (1997) pp. 219–222 and Thaddeus Metz, "The meaningful and the worthwhile: Clarifying the relationships," *The Philosophical Forum*, vol. 43, no. 4 (2012), pp. 443–446.

²¹ Other alternative values are more communitarian, reflecting the value of lives to a wider group or social unit, and may be a good fit for at least some of the goods listed by Parfit. These values are associated with critics of utilitarianism, such as those who support an ethics of care (see Virginia Held *The ethics of care: Personal, political, and global* (Oxford: OUP, 2006) p. 63), humanity/Ubuntu (see Thaddeus Metz, "Toward an African Moral Theory," *Journal of Political Philosophy* vol. 15, no. 3 (2007) pp. 335–337) or egalitarianism (Martin O'Neill, "What should egalitarians believe?," *Philosophy & Public Affairs* vol. 36, no. 2 (2008), p. 126). On many such views, love, friendship, bringing up children and acting rightly, may not be better for anyone in particular than other sources of welfare, but they create bonds of care, trust and common humanity and promote the good of the collective group constituted by these bonds. Hence, the value of such things is impersonal and cannot be reduced to individual welfare.

²² By the 'all things considered' value of a life, I mean the value it contributes to a population that contains it - leaving aside any distributional considerations such as whether it makes that population more or less equal.

2 The Problem with Perfectionism

As we have seen, perfectionism allows that many things, apart from the best things in life, determine the, all things considered, value of a life. These include lesser goods, such as everyday pleasures and achievements, neutral states, such as unconsciousness, and bad things that make lives go worse, such as suffering and frustration. Individual lives can be made up of many combinations of these things and are assigned an overall ‘welfare level’ that indicates the total value of this combination for the person living that life.

Nevertheless, Perfectionism claims that:

- c. Only lives that contain the best things in life can have a ‘very high Quality of Life’ and
- d. The best things in life have more than just welfare value, and a sufficient quantity of these things give a life containing them an, all things considered, value that is lexically superior to that of any amount of welfare on its own.

The first of these claims seems intuitively plausible. At least, it is hard to believe that someone who enjoys none of these things and is devoid of all meaning and other impersonal values may have a very high quality of life, even if their life contained many minor pleasures and achievements spread over time. While there are many ways of aggregating the value of temporal welfare within a person’s life, it is not hard to accept that one of the criteria for a life to be very high quality is that it must contain at least some of the best things in life.

However, these claims do not give us sufficient reason to find the Repugnant Conclusion very hard to accept, nor do they allow us to avoid it. This is because they only tell us about the composition of high-quality lives. Therefore, whilst perfectionism implies that lives with a very high Quality of Life have a value significantly greater than lives that lack any of the best things in life, it does not rule out some lives that are barely worth living also containing the best things in life as well.

For instance, lives that are barely worth living may contain some the best things in life if they contain only a few of these things, also contain some bad things, such as suffering and frustration, and are otherwise dull or short. These factors each serve to reduce the welfare level of such a life so that it could have very little value for the person living it – a life that is barely worth living. However, because each of these low welfare lives contains some of the best things in life, there must be some number of such lives that, together, would contain more of the best things in life than many lives with a very high Quality of Life. Therefore, if we value the total quantity of the best things in life in a population, we should be willing to accept that an outcome containing just these lives, though barely worth living, would be better than one containing a very large number of very high-quality lives. This implies the Repugnant Conclusion.²³

²³ See Jasper Ryberg “The Repugnant Conclusion and Worthwhile Living.” In Jasper Ryberg and Torbjörn Tännsjö. Eds., *The Repugnant Conclusion. Essays on Population Ethics* (Dordrecht: Kluwer Academic Publishers, 2004) p. 251.

It might be responded that this only follows if we valued the total quantity of the best things in life but that there are other ways we could aggregate the value of these goods across lives. For instance, Thomas Hurka has argued that we have a special reason to value the average, rather than the total, quantity of the best things in life, both across time and across lives. He points out that many of us feel that when a person's career has been intensely good, then prolonging it for the sake of a few more good things at the cost of lowering its overall quality reduces its value. He advocates extending this attitude "to human lives and then to all human history" – an extension he believes to be "appealing in itself".²⁴ If it can sometimes be worse to have more good things in one's life, even if this does not involve the loss of any of the best things in life, then we must value something other than the total quantity of the best things in life, for instance, their average quantity across lives and over time.²⁵

However, this response fails because we cannot rule out the possibility that some barely worth living lives may actually contain *more* of the best things in life than some very high-quality lives. This would be the case if they contained many of the best things in life, but also a great many bad things. No matter how much a particular person enjoyed the best things in life, their life might still be barely worth living.

Consider the biblical character Job. For the first part of Job's life, he had an exceedingly high quality of life, enjoying all the best things in life. However, at a certain point, his fortunes changed and Job's life became exceedingly bad. If the suffering and frustration of this part of Job's life went on for long enough, then its value to him would clearly fall until his life eventually became barely worth living (or even 'worth not living'). Job made no mistake when, in the midst of his suffering, he cursed the day on which he was born, even though had he not been born Job would have missed out on enjoying all the best things in life. The value of these things for Job was cancelled out by the badness of his later life. Nor is this the only way in which a life may be barely worth living and still contain some of the best things in life. For instance, Parfit mentioned similar cases, which he called 'roller coaster' lives, although he declined to comment on their value.^{26,27}

Denying that any life that contains some of the best things in life could be 'barely worth living' would imply that there is some level, below which a life that contained

²⁴ Hurka op. cit. p. 71.

²⁵ Hurka proposes, and indeed prefers, other alternatives for aggregating the value of perfectionist goods, including the less extreme claim that we should value both the average and the total quantity of these goods and the more extreme claim that we should value only the maximum quantity of these goods in any individual life or at any moment in time. However, all of these aggregation methods are equally vulnerable to the response I propose here.

²⁶ See Parfit (1986) op. cit. pp. 163–164 and (2016) op. cit. p. 218.

²⁷ There is some indication that Parfit found roller coaster lives more valuable than other kinds of low welfare life and that, despite accepting that they are 'barely worth living', he believed that if the Z population were to consist of these lives, then the Repugnant Conclusion would no longer be repugnant or very hard to accept (see Parfit (1986) op. cit. p. 19 – he also implied that this was the case in personal correspondence). I will not consider this view further here except to note that, as I argue in Sect. 4, it seems to have the troubling implication that lives that are bad for the people living them can still make an outcome better.

any of these things could not fall. However, if this were the case then even very bad features of some lives, such as long periods of intense suffering, might be morally irrelevant. This would be so when they were contained in the same life as many other bad things and some of the best things in life. In this case, since further bad features could not reduce the value of these lives, and hence their effect on the value of an outcome, beyond a certain level, they would have little relevance to our moral decision making.²⁸

For instance, this view would imply that it might be better if many people like Job, who enjoyed some of the best things in life but also faced many bad things, were to endure even more bad things than if a single person forwent a few additional ordinary good things. Since the Job-like people are already close to the lowest possible welfare level for somebody enjoying many of the best things in life, any amount of additional suffering for these people would reduce their welfare to a much lesser extent, if at all, than the absence of these few good things. Hence, these bad things would only have a minor impact on the value of the outcome as a whole.

I find this implication unacceptable and therefore reject the idea that lives that enjoy any of the best things in life could not be barely worth living. However, if this is so, then we can imagine a version of the Repugnant Conclusion where all the people in population A enjoys fewer of the best things in life than the people in population Z, even though these A people have a far higher Quality of Life than the Z people. No matter how we aggregate the value of the best things in life across these lives, therefore, perfectionism may imply the Repugnant Conclusion. Hence, we must conclude that perfectionism on its own is not a sufficient response to the Repugnant Conclusion.

3 Perfectionism and the Quality of Life

So far, I have presented two ways in which lives that are barely worth living, at a very low welfare level, might nevertheless contain the best things in life, and shown how these cases undermine perfectionism's sufficiency as a response to the problem of the Repugnant Conclusion. In both these cases, the lives at very low positive welfare levels had to contain bad things, such as suffering and frustration, suggesting that if a life is below a certain welfare level and contains some of the best things in life, then it must necessarily contain some bad things as well.²⁹

²⁸ This would violate what Tyler Cowen calls the 'non-vanishing value axiom', that for any one value, there exists no distribution of other values such that the addition of any amount of this value would not make up for the removal of any finite amount of some other value (see Tyler Cowen, "What Do We Learn from the Repugnant Conclusion?," *Ethics* vol. 106, no. 6 (1996), p. 759). Even if we do not think that *all* values are non-vanishing in this way, the badness of individual suffering surely *should* be non-vanishing.

²⁹ In earlier work, Parfit considered possible cases in which people live very low welfare lives whilst enjoying some of the best things in life and no bad things. In such cases, "The people in Z do each, once in their lives, have or engage in one of the best experiences or activities. But all of the rest is Muzak and Potatoes" (Parfit (1986) op. cit. p. 19). For the reasons I set out below, I do not find such cases credible.

In recent work, Parfit ceased to make use of such cases and stated that a life containing even a few of the best things in life and no suffering "could not be called barely worth living" (Parfit (2016) op. cit. p. 118).

There is another reason to believe that any life that is barely worth living which contained some of the best things in life must also contain bad things. The best things in life, as conceived by Parfit and others, are something more than momentary experiences or natural capacities that anyone might have. They are the results of a fulfilling and fulfilled life, a point that is only strengthened if we conceive of the special value of these things as being something like meaning. It would, therefore, be impossible for anyone to enjoy some of the best things in life without also enjoying many other good things. For instance, whilst young babies might enjoy listening to music by Mozart and Bach, this enjoyment does not constitute one of the best things in life. Listening to such music only becomes truly meaningful or fulfilling once people have developed their aesthetic sense and the capacities of appreciation and interpretation. Such development involves many experiences and activities that are good in themselves and make life worth living. Other examples of the best things in life, such as scientific discovery, true friendship or child rearing similarly require that a person's life contain many other good things, and would not possess the same meaning, or other impersonal value, if they merely 'happened' to us. Therefore, a person who enjoys any of these best things in life, even to a minimal extent, must have a life with features that would, *ceteris paribus*, already make it valuable for that person and hence must be at a significantly positive welfare level.³⁰ They could not then fall below this level unless their life also contained bad things, such as suffering or frustration.³¹

This fact about the composition of barely worth living lives allows us to develop another view that more successfully responds to the problem of the Repugnant Conclusion. On this view, not only do the best things in life have a, superior, value that is not shared by ordinary good things, but bad things also have a superior (dis)value that is also asymmetrical with the value of ordinary good things, such as minor pleasures and achievements, yet symmetrical with that of the best things in life.

To illustrate this view, let us imagine the combination of one ordinary good thing (a quiet walk) and one bad thing (a mild headache). These might be such that the addition of both of them to any given life has no net effect on its welfare level, i.e. together they make that life go neither better nor worse for the person who lives it. However, if the value of bad things is not symmetrical to that of good things, it may be better, all things considered, to cure someone suffering from the headache rather than providing someone with the opportunity to take the walk. This position has been advanced, and to some extent defended, by Jamie Mayerfeld.³²

Some philosophers, including Mayerfeld and Parfit, argue that such claims about the asymmetric value of pleasure and pain cannot be true at the fundamentally evaluative level, but rather reflect a failure to evaluate these things on the

³⁰ This is not the only view that would support this conclusion. One might simply believe that the best things in life were so enjoyable that they would make any life that contained them much better for the person living it.

³¹ See Douglas Portmore, "Does the total principle have any repugnant implications?," *Ratio* vol. 12, no. 1 (1999), p. 86.

³² Jamie Mayerfeld, *Suffering and Moral Responsibility* (Oxford: OUP, 1999), pp. 128–131.

correct scale. For instance, they argue that the apparent asymmetry in these cases is a result of the fact that we are actually measuring something like the degree to which pleasure or pain appears good or bad to a person at the time they are felt, which may not be directly related to their contribution to a person's overall welfare level. If we correctly evaluated pain and pleasure relative to their actual contribution to people's welfare, these philosophers argue, then there could no longer be an asymmetry between the value of good and bad things in that life; although we may still have greater reason to help those who are suffering, for instance if we accept prioritarianism.

However, whilst this may be true for many views that incorporate such an asymmetrical value, it is not true that all value asymmetries face this kind of objection. It may be that, even if good and bad things really do contribute equally to the welfare level of a life, the badness of bad things is qualitatively different to the goodness of good things, so that the bad things contribute more to the value of the life, all things considered, because they are both personally and impersonally bad. On one account, this qualitative difference is a matter of justice. For instance, bad things may be the source of a moral complaint, where the existence of such complaints is bad even though they do not make a person's life go worse. Enduring suffering and frustration may, therefore, reduce the value of one's life in a similar way to being the worst-off person in an unequal society or failing to get what one deserves.³³ On another account, this qualitative difference would be more purely axiological, because bad things, like the best things in life, possess distinct impersonal (dis)value. For instance, suffering and frustration may actively destroy meaning or may carry additional, purely negative, values such as alienation or anti-meaning.^{34,35}

To illustrate this, let us return to our previous example and consider a case in which someone is willing to take a quiet walk knowing that she will get a mild headache afterwards because she knows that the net effect of these things on her welfare will not be negative. The view that I have just sketched implies that there could remain a negative balance of impersonal value if she does this, so that her act might make the world worse, though it did not make her life go worse for her and affected nobody else. How can this be? One reason is that anyone in a state of suffering has a

³³ See Scanlon *What we owe to each other* (Cambridge Ma: Harvard University Press, 2000), pp. 226–227 and O'Neil op. cit. pp. 121–122.

³⁴ Stephen M. Campbell and Sven Nyholm, "Anti-meaning and why it matters," *Journal of the American Philosophical Association*, vol. 1 no.4 (2015) 694–711.

³⁵ I have stated that these additional values apply to all bad things in a life and continue to claim this in the rest of this paper. However, certain very minor bad things may plausibly make no qualitative difference to an otherwise good life. These may include, negative experiences that form an intrinsic part of positive experiences (like masochistic pain), frustrated desires that are substituted by satisfied desires (like not being able to eat strawberry ice cream when I can eat raspberry, which I like just as much) or bad things that fail to dominate my life for even a moment (like a single isolated hangnail or mosquito bite, in the midst of an otherwise good life). This makes little difference to my theory however, since if we made these bad things worse or more numerous in a person's life, or if they formed part of an otherwise neutral life or one with many other bad things in it, then they would make a qualitative difference to that life and possess these additional values. As with the nature of impersonal values that the best things in life possess, I believe that this topic warrants further study.

general claim to have this suffering alleviated, and that a failure to satisfy this claim represents an injustice that should be viewed as making this outcome worse. Whilst we might believe that because this pain was self-inflicted, such a claim would be weak and its non-satisfaction less bad, this still seems to be a negative feature of this outcome that makes it qualitatively different to that in which this person stayed at home. Another reason is that it is wrong only to consider the contribution that this person's headache made to their welfare as a whole, since suffering also affects their life in other ways, such as precluding their realization of aesthetic and moral value, and that suffering is not only bad for a person as a whole but is especially bad for the person stage who must endure that suffering. For these reasons, we may well be acting hastily if we jump straight from the observation that a person, correctly, judges that a certain bundle of goods and bads will be good for them to the judgement that it could not make the world worse.

These kinds of consideration have motivated several discussions about the apparently asymmetrical value of good and bad things in people's lives.³⁶ However, for this paper it is enough to note that the kind of asymmetry I am trying to defend is one where both good and bad things have symmetrical effects on the welfare value of lives, but bad things have an asymmetrical effect on the, all things considered, value of lives relative to ordinary good things, because of the qualitative differences between them. This asymmetry is such that:

- e. As with the best things in life, bad things have more than just welfare value, and a sufficient quantity of bad things has a, negative, value that is lexically superior to the, positive, value of any quantity of welfare on its own and
- f. For any quantity of the best things in life, some, sufficiently large, quantity of bad things would have a greater (negative) value, whilst for any quantity of bad things, some, sufficiently large, quantity of the best things in life would have a greater (positive) value.³⁷

Accepting both perfectionism and this asymmetrical value for bad things allows us to formulate a sufficient response to the problem of the Repugnant Conclusion. Just as we believed that no quantity of additional welfare made a population void of the best things in life better than one in which these things are enjoyed to a high degree, so we might believe that no quantity of additional welfare could make a population that is free from bad things better than one that is not. However, we might

³⁶ See Karl Popper, *The open society and its enemies: volume 1* (London: Routledge, 1995), p. 317, Mayerfeld op. cit. pp. 145–158, Clark Wolf, "O Repugnance, Where Is Thy Sting?" In Jasper Ryberg and Torbjörn Tännsjö. Eds., *The Repugnant Conclusion. Essays on Population Ethics* (Dordrecht: Kluwer Academic Publishers, 2004) pp. 71–72 and Torbjörn Tännsjö, "Utilitarianism or Prioritarianism?," *Utilitas*, vol. 27, no. 2, pp. 242–245.

³⁷ This symmetry between the best things in life and bad things seems, to me, to chime with one of Parfit's very last statements about the value of lives, that "Some of our successors might live lives and create worlds that, though failing to justify past suffering, would give us all, including some of those who have suffered, reasons to be glad that the Universe exists" (Derek Parfit, *On What Matters Volume III* (Oxford: OUP, 2017) p. 437).

well believe that even large amounts of suffering, frustration and other bad things could be worthwhile if they were necessary for there to be an even larger amount of the best things in life and that it might similarly be worth sacrificing some of the best things in life if it were the only way to remove even greater amounts of bad things from the world.

This response explains the repugnance of the Repugnant Conclusion in all the cases I have mentioned so far and should allow us to avoid it in all other cases as well. This is because, according to the first perfectionist claim, c, a population of people who all have a 'very high quality of life' will necessarily contain many of the best things in life. On the other hand, because its lives are below the welfare level of a life containing any of the best things in life and no bad things, any population where everyone has a life that is barely worth living will either contain lives without any of the best things in life (and thus have a lower value according to the second perfectionist claim, d) or lives with more bad things than the best things in life (and thus have a lower value according to the second claim about the asymmetrical value of bad things, f). Together therefore, these three claims justify the view that, even though the Z population contains far more of what makes life worth living than the A population, the qualitative differences between the lives in each population always mean that the Z population will be no better than the A population, all things considered, i.e. they justify rejecting the Repugnant Conclusion.³⁸

Therefore, even though perfectionism on its own is not a sufficient response to the problem of the Repugnant Conclusion, this combined view about Quality of Life, incorporating both perfectionism and an asymmetrical value for bad things, i.e. principles c, d and f, may be so.

Parfit suggested another objection to perfectionism that gives us additional reasons to accept a combined view like this. He argues that, on its own, perfectionism may have implications that are unacceptably elitist and show a disregard for those who are worst off. For instance, the belief that a sufficient quantity of the best things in life have a value that is lexically superior to that of any amount of welfare on its own, implies the very unattractive view that "the prevention of great suffering can be ranked wholly below the preservation of creation of the best things in life", and hence the 'Elitist Conclusion' that it is always better to allow a small group of people to enjoy the best things in life than to save a far greater number from experiencing terrible suffering.³⁹ Whilst Parfit declined to propose any response to this objection, it seems clear that combining perfectionism with the claim that 'there is no quantity of the best things in life such that a sufficiently large quantity of bad things could not have a greater, negative, value', would be a direct and compelling means of avoiding this objectionable implication.

³⁸ Note that claim e, concerning the asymmetrical value of bad things, is not strictly necessary for this explanation. Some may choose to reject this claim and doing so would be compatible with the other arguments of this paper. Nevertheless, it seems to me that f is most plausible when we also accept e, hence my reason for including it above.

³⁹ Parfit (1986), op. cit. p. 20.

4 Is my view ‘Ridiculous’?

On the other hand, Parfit has identified an important objection to views that incorporate an asymmetrical value for bad things. According to:

*The Ridiculous Conclusion: Compared with the existence of many people who would all have some high quality of life, there is some much larger number of people whose existence would be worse, even though these people would all have a much higher quality of life.*⁴⁰

Parfit devised this conclusion as an objection to the view that there might be an upper limit on the goodness of positive welfare in a population, but not on the badness of negative welfare. However, it is also implied by other views that include an asymmetrical value for bad things, such as the one I am proposing here. In all such cases, the conclusion is implied when “this much larger population would be worse, because in each of these lives there would be some intense suffering.”⁴¹

Parfit did not elaborate on this conclusion in his published work, and it has received little attention in the literature since. In responding to this objection I, therefore, intend to take a thorough approach and note that the ridiculousness of the ridiculous conclusion seems to rest on the conjunction of three different claims: first, lives that are better for the people living them should not be worse than those that are worse for the people living them; second, the mere addition of good lives to a population should not make it worse and third, the strength of these two claims increases when the lives under consideration are of a high quality. Since all three of these claims appear compelling and each of them raises its own issues with the theory I have sketched here, I will address each of them in turn.⁴²

Let me begin by addressing the first of these claims. It is true that, if the value of bad things is not symmetrical to that of good things, then lives that are better for the people living them, i.e. that are at a higher welfare level, can be worse, all things considered, than those that are worse for the people living them, a fact that we have already discussed in previous sections. Consider the following populations⁴³:

⁴⁰ I have adapted Parfit’s presentation of the Ridiculous Conclusion from *Reasons and Persons* to make it match his most recent statements of the Repugnant Conclusion. Parfit’s original version was as follows: “if there were ten billion people living, all with a quality of life about that of the average quality of lives lived by the world’s present population, there must be some much larger imaginable population whose existence would be worse, even though all of its members would have a very much higher quality of life” (Parfit (1984) op. cit. p. 407).

⁴¹ Parfit (1984), op. cit. p. 407.

⁴² In response to an early draft of this paper, Parfit accepted this interpretation of his conclusion and endorsed this approach to dealing with it.

⁴³ In this, and later, examples, I apply numbers to the welfare level of lives and the quantity of good and bad things within them for illustrative purposes. These numbers indicate ordinal relations of value for the person whose life contains them. For instance, 15 units of bad things indicates some amount of bad things that would make a person’s life worse, for that person, than the absence of 10 units of good things, whilst a life at welfare level 100 implies a life that is better, again for the person living it, than a life at welfare level 95.

Population P) n people at level 100, consisting of 110 units of good things and 10 units of bad things.

Population Q) n people at level 95, consisting of 95 units of good things, and no bad things.

If the value of good things and bad things is asymmetrical then, in cases like this, it is perfectly possible to believe that the addition of 15 units of good things and 10 units of bad things to each life makes population P worse than population Q, all things considered. This violates what seems like a fundamental principle of population ethics, namely that a perfectly equal population at a higher welfare level is always better than a perfectly equal population of the same size at a lower welfare level.⁴⁴ However, although this implication may seem disturbing, it merely highlights what it means to hold any view on which the value of a life is not the same as its welfare value. To argue that higher welfare lives are always better than lower welfare lives in this case simply begs the question against asymmetrical value views based on the impersonal value of bad things.

Furthermore, perfectionism has a very similar implication. As we have already noted, lives at lower welfare levels can contain more of the best things in life than lives at higher welfare levels. If we do not take into account the special badness of bad things, we face the following conclusion:

Population R) n people at level 100, consisting of 100 units of good things, but none of the best things in life.

Population S) n people at level 95, consisting of 110 units of good things (including some of the best things in life) and 15 units of bad things.

According to the kind of perfectionism proposed by Parfit, in which “even if some change brings a great net benefit ... it is a change for the worse if it involves the loss of one of the best things in life.” population S is better than population R. This is so even though everyone in population R is at a higher welfare level and everyone in population S endures some bad things. Indeed, in his most recent work, Parfit specifically argued that lower welfare lives could be better than higher welfare lives if they enjoy more of the best things in life.⁴⁵ This suggests that no Perfectionist could reject my view as ridiculous merely for having this implication.

Let me next turn to the claim that the mere addition of good lives to a population cannot make it worse. Again, my view would not be consistent with this claim because the bad things in a life may mean that it has a negative value, even if it is good for the person living it. It follows that not only does this view violate principle b, that ‘the value of all lives can be represented on a finite common scale’, but, in some cases, it may also violate principle a, the ‘simple view,’ as well.

This may present a bigger problem for my view, since Parfit appeared committed to upholding the simple view. For one thing, he specifically endorses it in his

⁴⁴ See Gustaf Arrhenius, “An impossibility theorem for welfarist axiologies,” *Economics & Philosophy* vol. 16, no. 2 (2000), p. 61.

⁴⁵ See Parfit (2016) op. cit. p. 117.

work.⁴⁶ For another, he also defended the related principle that even if the addition of people with lives worth living lowers the average welfare level of a population and makes its distribution of welfare much less equal this would not be bad.⁴⁷

Yet, this objection may also be question begging. After all, part of the justification for the simple view seems to be the claim that, if a life is good for the person living it, then its mere presence cannot make a population worse. However, Parfit does not appear to accept the view that benefiting a person never makes an outcome worse. For instance, if this benefit meant enjoying fewer of the best things in life then he is willing to accept that this would be bad. Yet, if we accept that benefiting someone could make an outcome worse, then why is it ridiculous to believe that it is sometimes worse for a person to exist with a life that is good for them than for them to have no life at all?

A Perfectionist might respond that whilst their view does distinguish between different kinds of value, this is not the only justification for principles like the Simple View, and that perfectionism itself is consistent with this view, whatever it may imply about some of the justifications for it. This is because enjoying the best things in life can only increase the value of a life beyond that of its welfare level; it can never reduce it.

However, perfectionism has another implication that may be even more troubling, that it can be good to bring into existence people whose lives are bad for them. Since, according to perfectionism, lives at very low welfare levels can still have significant positive value if they contain the best things in life, and since the addition of more of the best things in life is a change for the better, even if it involves a great net cost to those affected, some lives with negative welfare could have positive value. Parfit however, appeared even more committed to the view that lives that are bad for the people living them cannot make an outcome better than he was to either the simple view or the mere addition principle.⁴⁸ However, note that if we also accept an asymmetrical value for bad things, our combined view could avoid this implication, since bad lives would always contain many more bad things than good ones – let alone the best things in life. For this reason, I conclude that no Perfectionist could reject my view as ridiculous merely for implying that the addition of ‘good’ lives can sometimes make an outcome worse.

Finally, therefore, let us consider the claim that increasing the welfare level of lives or adding additional good lives must at least make an outcome better where these lives are all of a high quality. I find this claim far more compelling. However, note that in both of the above cases, the reasons why this combined view sometimes appears to imply the ridiculous conclusion is precisely because it takes into consideration a person’s quality of life as well as their welfare level. Lives at higher welfare levels can only be worse if they are of lower quality. Similarly, lives that are good

⁴⁶ See Parfit (2016) op. cit. p. 110.

⁴⁷ See Parfit (1984) op. cit. p. 420.

⁴⁸ I say this because he appeared to find cases where the existence of people with bad lives is supposed to make an outcome less bad, such as his Hell 3 or Gustaf Arrhenius’s Sadistic Conclusion, even harder to accept than the Repugnant Conclusion (see Parfit (1984) op. cit. p. 422 and (2016) op. cit. p. 111).

for the person living them can only make an outcome worse if they have a very low quality of life. When it is stipulated, as in the Ridiculous Conclusion, that a person has a higher quality of life, then neither of these two, somewhat troubling, implications follow.

This combined view about the qualitative value of both bad things and the best things in life, therefore, does not imply Parfit's ridiculous conclusion, even though it does appear to violate two of the claims that support it under certain conditions. Furthermore, in violating these two claims, it does so in ways that nobody who accepts perfectionism should find objectionable. It will violate the first of these claims less often than either perfectionism or an asymmetrical value view would on their own, and, unlike perfectionism on its own, it can avoid the implication that lives that are bad for the people living them make outcomes better.

5 Quality of Life and Population Ethics

In previous sections, I have shown how a combined view including perfectionism and an asymmetrical value for bad things gives us both a defensible means of avoiding the Repugnant Conclusion and an argument for rejecting some of the principles that imply it. However, these are not the only intuitively plausible principles that imply the Repugnant Conclusion. In this section, I will show how this combined view can also give us reason to reject other commonly held principles and allow us to avoid the paradoxes they create in a compelling way.

Let us consider the Mere Addition Paradox.⁴⁹

This paradox is that, for the following three populations, A+ is better than A and Z is better than A+, but Z is no better than A:

Population A) a very large number of people with a very high quality of life

Population A+) the same number of people with at the same very high quality of life and a much larger number of people with lives that are only barely worth living

Population Z) the same number of people as in population A+, all of whom have lives that are also only barely worth living but that are slightly better than the worst-off lives in population A+, so that this population has the greatest total quantity of welfare.

These populations are illustrated below (Fig. 2):

Let us assume that population A and population Z are such that to claim that Z was better than A would imply the Repugnant Conclusion. In this paper, I have argued that the best explanation for why Z is no better than A is that, whilst the lives in A (and hence the best-off lives in A+) contain many of the best things in life, the lives in Z (and hence the worst-off lives in A+) contain either none of the best things in life or they contain many bad things. It follows that there is a significant

⁴⁹ See Parfit (1984) op. cit. pp. 425–430, Huemer (2008) op. cit. pp. 901–903 and Arrhenius (forthcoming) op. cit. p. 310–314.

qualitative difference between the lives in A and the lives in Z and that the value of this difference is greater than any additional quantity of welfare in Z.

It is claimed that A+ must be better than A, since A+ can be formed from A by Mere Addition. This is the addition of “extra people 1) who have lives worth living, 2) who affect no one else, and 3) whose existence does not involve social injustice”.⁵⁰ As we have seen, according to the simple view, each of these additional lives should make the world better. However, if we accept an asymmetrical value for bad things, then it is possible that mere addition can make a population worse. This will happen if the extra people endure many bad things, such as suffering or frustration. According to an asymmetrical value view, whilst such lives might be good for the people living them, the special badness of the bad things they contain makes these lives bad, all things considered. It would, therefore, follow that A+ is only better than A if its lives do not contain many more bad things.

Similarly, it is claimed that Z must be better than A+, since Z contains more welfare in total, a higher average welfare level and a more equal distribution of welfare than A+.⁵¹ However, according to perfectionism and asymmetrical value views, such considerations will not imply that Z is better than A if Z contains many fewer of the best things in life or many more bad things. Even if one should prefer an equal distribution of welfare on its own, this may not be sufficient to warrant either removing some of the best things in life from, or adding bad things to, a population.

It follows that, according to a combined view like the one I have described here, we may find that whilst it is possible for A+ to be better than A and Z to be better than A+, the conditions under which these two judgements might be true are such that, for any actual triplet of populations, either A+ will be no better than A, because it contains more bad things, or Z will be no better than A+, either because it contains more bad things or fewer of the best things in life. However, which of these possibilities is the case is underdetermined by the information provided in this example about peoples' welfare level and quality of life. Hence, the relative value of these populations will depend on what sort of lives they contain. This leaves us with two possibilities:

Firstly, If A+ were better than A, then the worse off lives in A+ must not contain many bad things. However, Z is only better than A+ if it does not contain either many fewer of the best things in life or many more bad things. Since the best-off lives in A+, like the lives in A, are at a very high quality of life, they must contain many of the best things in life. However, since the lives in Z are only barely worth living, they will contain none of the best things in life unless they also contain many bad things. If the lives in Z do not contain any of the best things in life, then Z will be no better than A+. However, if the lives in Z do contain many bad things, then they must also be no better than the lives in A+, since the A+ lives do not contain many bad things. Hence, if A+ is better than A, then Z will be no better than A+, despite having a greater total quantity of welfare.

⁵⁰ See Parfit (1984) *op. cit.* p. 420.

⁵¹ *Ibid* p. 421.

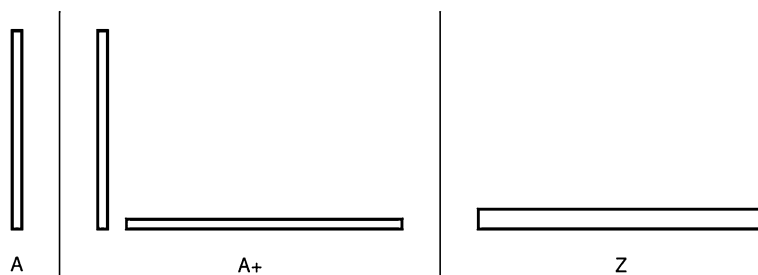


Fig. 2 The Mere Addition Paradox

On the other hand, if Z were to be better than A+, then Z must not contain many fewer of the best things in life or many more bad things than A+. However, if the lives in Z, which are all barely worth living, are not to contain many fewer of the best things in life than the better off, high-quality, lives in A+, then the Z lives must also contain many bad things. It would follow that the worse off lives in A+ would also have to contain many bad things, and the addition of these lives would make this population no better, even though they are all worth living. Hence, if Z is better than A+, then A+ will be no better than A, despite it being the result of mere addition.⁵²

⁵² Note that whilst my theory allows for this solution it does not, as yet, prescribe it, because it does not specify key details about how to evaluate the various components of Quality of Life. For instance, a reviewer of this paper suggested that my theory might not have this result for a case in which the lives of the worse-off group in A+ could be filled with both the best things in life and bad things, but with a net-positive balance of the best things so as to make A+ better than A, and in which Z was obtained from A+ by redistributing some of the bad things in each of these lives to the better-off group, lowering their quality of life, whilst also reducing the total quantity of bad things in Z compared to A+ and even adding some more of the best things in life amongst the whole population. This is because, if the worse off group in A+ is large enough then, it should be possible to achieve this transformation without making any of the lives in Z more than 'barely worth living' and yet making Z better than A+ and hence better than A. However, this assumes a roughly linear trade-off between changes in the impersonal and personal value of the best things in life and bad things. If we do not assume this, then we need not accept that lives that are full of both the best things in life and of bad things, but with a net-positive balance of the best things in life, must be barely worth living. As I described in the previous section, it seems to me that it would be worse for a theory to imply that some lives that were bad for the people living them would make the world better than for it to imply that some lives that are good for the people living them might make the world worse. For this reason, and given the fact that welfare level and Quality of Life are only loosely connected, it seems reasonable to say that where a life is full of both bad things and the best things in life, then a barely worth living life will be one in which the balance of impersonal values is negative and, that a positive balance would lead to lives that are more than barely worth living. If this were so, then the lives described in this objection will be different to those of the worse off group in A+. For a life to be only barely worth living and yet contain a net positive balance of the best things in life, on this view, it would need to contain very few of these things, so that if some of these bad things were taken away, or some more of the best things in life were added, then the balance of bad things and the best things in life would change considerably, and these lives would cease to be barely worth living. I am grateful for this reviewer's comment and acknowledge that developing a theory of Quality of Life that supports all the judgements I have here described without having other undesirable implications remains challenging.

Therefore, at least one of the steps in the mere addition paradox fails to apply. However, we have no way of knowing which step is at fault, since we have no information about the relative amounts of bad things or the best things in life in each of these populations. All we know is that there may be no allocation of these goods across the lives in these populations such as to replicate this paradox under the combined view that I have proposed here. We also have good reason to assume, at least *ceteris paribus*, that the intuitions that lead us to accept each step of this paradox are correct, since they are expressed in terms of the value of welfare and only conflict in a few extreme cases involving big differences in quality of life, such as this one. Hence, the sense of paradox that these populations evoke is retained, without implying that we have any genuinely non-transitive judgements about the value of the populations involved, whilst the paradox itself dissolves once we appreciate how the composition of the lives in each of these populations can make a difference to our judgments about their value.

On the other hand, if we only accepted perfectionism, then this would lead us to the conclusion that the only possible solution to the mere addition paradox is that population Z is no better than population A+, because Z contained fewer, if any, of the best things in life. Not only is this untrue, since as I have argued Z could contain *more* of the best things in life than A+, but it is also counterintuitive in that it denies the sense of paradox we feel when considering these populations. Perfectionism, on its own, leaves no space for our intuitive sense that our judgment about the relative value of A+ and Z is no less problematic than our judgment about the relative value of A and A+, yet for me, it has always seemed that this is so. For this reason, I find the solution of my combined view more attractive than that of perfectionism on its own.

This is not the only paradox in population ethics.⁵³ My combined view can, I believe, offer similar responses to these other paradoxes as well. However, I shall leave discussion of such responses for further work.

6 Conclusion

In this paper, I have considered some possible responses to the problem of the Repugnant Conclusion and its associated paradoxes. It is often assumed that we can explain and avoid this conclusion by an appeal to the perfectionist claim that we ought to value certain goods that are contained in lives of very high quality, the best things in life, as lexically superior to other kinds of welfare. However, I have shown that this is not so – we can neither explain nor avoid the Repugnant Conclusion by appealing to perfectionism alone. This is because it is possible that a large enough number of lives that are barely worth living will nevertheless contain more of the best things in life than very high-quality lives. Instead, I have suggested that only a combined view, which incorporates both perfectionism and an asymmetrical value

⁵³ See Cowen (1996) *op. cit.* pp. 757–760, Rachels (2004) *op. cit.* pp. 163–168, Parfit (2016) *op. cit.* pp. 119–127 and Arrhenius (Forthcoming) *op. cit.* pp. 315–357.

for bad things, could be a sufficient response to the problem of the Repugnant Conclusion. I have defended such a view from the objection that it implied the ‘Ridiculous Conclusion’ and argued that in this respect it performs at least as well as, and in some ways better than, perfectionism. Finally, I have shown how the way in which this combined view responds to the problem of the Repugnant Conclusion allows us to avoid the Mere Addition Paradox and other problems in population ethics.

Acknowledgements I am grateful for feedback on this paper from Derek Parfit, Gustaf Arrhenius, Tim Campbell, Patrick Kaczmarek, Partha Dasgupta, Hilary Greaves, Michael Otsuka, Alex Voorheve, audiences at the Institute for Philosophy’s Spring Graduate Conference 2015 and the Theoretical Population Ethics Conference at the University of Oxford and seven anonymous reviewers. This project was made possible through the support of a grant from Templeton World Charity Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of Templeton World Charity Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.